

**Working Group:**

**Instrument name:**

**Domain:**

**Date of submission:**

Table of Contents

[Introduction 4](#_Toc75507615)

[Assembly of working group and protocol development 6](#_Toc75507616)

[1. Assemble working group 6](#_Toc75507617)

[2. Decide on methods protocol for Core Outcome Instrument Set selection 7](#_Toc75507618)

[2.1 Target PICOC of Instrument 7](#_Toc75507619)

[2.2 Methods to identify candidate instruments 8](#_Toc75507620)

[2.3 Description of candidate instrument 8](#_Toc75507621)

[2.4 Protocol methods plan 9](#_Toc75507622)

[3. Deliverable: Submit protocol using Instrument Selection Workbook to Technical Advisory Group [TAG] 10](#_Toc75507623)

[4. Review and approval of final protocol by TAG 10](#_Toc75507624)

[Review of evidence of instrument performance for existing or new instrument 11](#_Toc75507625)

[Part A: Domain match and Feasibility Assessment 11](#_Toc75507626)

[5. Obtain Working Group and others assessment of match with the target domain 11](#_Toc75507627)

[5. 1 Survey of working group members about the content and domain match 11](#_Toc75507628)

[5.2 Survey of patients and other key stakeholders about the content and domain match 13](#_Toc75507629)

[5.3 Review of Raw Data 13](#_Toc75507630)

[Review of Raw Data Form 14](#_Toc75507631)

[6. Obtain Working Group and others assessment of feasibility 15](#_Toc75507632)

[6.1 Survey of working group members about feasibility 16](#_Toc75507633)

[6.2 Survey of patients and other key stakeholders about feasibility 17](#_Toc75507634)

[7. Obtain working group decision based on synthesis of overall ratings of domain match and feasibility 18](#_Toc75507635)

[PART B: Review of the evidence on measurement properties 19](#_Toc75507636)

[8. Conduct literature search to find studies; create PRISMA diagram, fill in Summary of Measurement Properties (SOMP) Table with results of which measurement properties are assessed in each included study 19](#_Toc75507637)

[8.1. Design a search strategy for the literature search 20](#_Toc75507638)

[8.2 Documenting your actual search strategy 20](#_Toc75507639)

[8.3 Screening and selection 22](#_Toc75507640)

[8.4 Creating your PRISMA Flow diagram 24](#_Toc75507641)

[8.5 Creating your Summary of Measurement Property (SOMP) Table 25](#_Toc75507642)

[9. Check to see if each of the included studies has used good methods when assessing each measurement property using the COSMIN-OMERACT good methods check; add these findings into the SOMP Table by coloring cells Green, Amber, or Red. 27](#_Toc75507643)

[9.1 COSMIN-OMERACT Good Methods Check 27](#_Toc75507644)

[9.2 Fill in SOMP with results of Good Methods Check 32](#_Toc75507645)

[10. Conduct data extraction on those measurement properties that were assessed as green or amber Good Methods and complete measurement property tables for summary descriptions of the studies; fill in SOMP Table with assessment of the adequacy of results 32](#_Toc75507646)

[10.1 Conduct data extraction and summary tables for each measurement property 32](#_Toc75507647)

[10.2 Judging the PERFORMANCE of the instrument based on the results found in the studies. 32](#_Toc75507648)

[11. Conduct synthesis across evidence available for each measurement property; fill in rating (Green/Amber/Red/White) on SOMP Table 34](#_Toc75507649)

[12. Decide if any gaps exist in evidence of measurement properties. If gaps found, draft protocol for new studies to fill gaps. If no gaps, fill in SOMP Table with proposed level of endorsement 34](#_Toc75507650)

[12. 1 Addressing gaps in the evidence 34](#_Toc75507651)

[12.2 Synthesis to obtain proposed overall rating of the instrument 34](#_Toc75507652)

[12.3 Complete SOMP table with proposed level of endorsement 34](#_Toc75507653)

[13. Deliverable: Submit the Instrument Selection Workbook to TAG 36](#_Toc75507654)

[13.1 Assemble documentation for submission 36](#_Toc75507655)

[13.2 Draft and submit protocol for new measurement property study 36](#_Toc75507656)

[14. Receive final response from TAG 36](#_Toc75507657)

[15. If studies are needed to fill gaps, conduct new measurement property studies, submit to TAG for Good Methods check, add to body of evidence (SOMP) and go back to Step 12 37](#_Toc75507658)

[If no studies are needed, put X here: \_\_\_\_\_\_and move to Step 16 37](#_Toc75507659)

[16. Obtain final agreement on final report 37](#_Toc75507660)

[17. Set timeline for next review of instrument 37](#_Toc75507661)

[18. Ratification of level of endorsement by OMERACT Community and communication of results 37](#_Toc75507662)

[19. Implement communication and dissemination plan 38](#_Toc75507663)

[Acknowledgements. 39](#_Toc75507664)

[Appendix A: Sample survey questions for surveying patients/other respondents on domain match and feasibility 40](#_Toc75507665)

[Appendix B: Sample form to help design your search strategy 42](#_Toc75507666)

[Appendix C. Search Strategies to be used for measurement property studies. 44](#_Toc75507667)

# Introduction

The (Outcome Measures in Rheumatology) OMERACT Handbook group created this workbook with detailed descriptions of the steps involved in instrument selection and potential tools for the job. We provide text for search strategies and checklists to help working groups move through various steps in the OMERACT Filter 2.2 Instrument Selection Algorithm (see Algorithm steps in figure below).

This workbook is designed to be used alongside Chapter 5: Instrument Selection of the OMERACT Handbook. The Handbook describes the theory and detail behind each step while the workbook is focused on presenting the results of your work. At the beginning of the workbook we have outlined the individual steps template on the methods of your review of the existing evidence of instrument performance. We encourage you to register your literature review on public platforms such as PROSPERO. While this is not mandatory, working groups are encouraged to register their work plan to add to the transparency and facilitate the publication of the results in the future.

We hope that the accompanying workbooks and templates help with tracking the steps and organizing information for your own use in publications and in presentations back to the OMERACT Technical Advisory Group and OMERACT Community.

**FOR A MORE DETAILED, THEORETICAL DESCRIPTION OF THE MEASUREMENT PROPERTIES BEING STUDIED AND EXAMPLES OF WHAT WOULD BE ACCEPTABLE TYPES OF EVIDENCE**

**SEE CHAPTER 5, SECTION 8 OF THE OMERACT HANDBOOK**

On the next page is the OMERACT Instrument Selection Process Checklist.

***Please check off the steps you have completed prior to submitting this workbook to the Technical Advisory Group for review.***

|  |  |  |  |
| --- | --- | --- | --- |
| **OMERACT Master Checklist for Instrument Selection *Name of Instrument*:** | | |  |
| **Step #** | **OMERACT Instrument Selection Process Checklist Item** | Mark when complete |  |
| **Assembly of working group and protocol development** | | |
| 1 | Assemble working group | **○** |
| 2 | Decide on methods protocol for Core Outcome Instrument Set selection | **○** |
| 3 | **Deliverable**: Submit protocol using Instrument Selection Workbook to Technical Advisory Group [TAG] | **○** |
| 4 | Review and approval of final protocol by TAG | **○** |
| **Review of evidence of instrument performance for existing or new instrument** | | |
| ***Part A: Domain match and Feasibility assessment*** | | |
| 5 | Obtain Working Group and others assessment of match with the target domain | **○** |
| 6 | Obtain Working Group and others assessment of feasibility | **○** |
| 7 | Is the instrument a match with the domain AND feasible? Yes \_\_\_\_  if yes, continue with Part B of checklist below No \_\_\_\_  If no, set instrument aside (find new one or develop new one) | **○** |
| ***Part B: Review of evidence of performance of an instrument across key measurement properties*** | | |
| 8 | Conduct literature search; create PRISMA diagram; place articles of measurement properties in Summary of Measurement Properties (SOMP) Table | **○** |
| 9 | Conduct COSMIN-OMERACT Good Methods check, add findings into the SOMP Table | **○** |
| 10 | Conduct data extraction, create summary reporting tables, fill in SOMP Table with assessment of the adequacy of results | **○** |
| 11 | Conduct synthesis across evidence available for each measurement property | **○** |
| 12 | Decide if any gaps exist in evidence of measurement properties  If gaps found, draft protocol for new study to fill gaps If no gaps, finish the SOMP Table with proposed level of endorsement | **○** |
| **Initial submission to TAG: literature review findings & protocol for gaps** | | |
| 13 | **Deliverable**: Submit the Instrument Selection Workbook to TAG | **○** |
| 14 | Receive final response from TAG | **○** |
| 15 | If studies are needed to fill gaps, conduct new measurement property studies, submit to TAG for Good Methods check, add to body of evidence (SOMP) and go back to Step 12 If no studies are needed, put X here: \_\_\_\_\_\_and move to Step 16 | **○** |
| **Final submission to TAG for approval** | | |
| 16 | Obtain agreement on final report | **○** |
| 17 | Set timeline for next review of instrument | **○** |
| **Ratification of level of endorsement by OMERACT Community and communication of results** | | |
| 18 | Ratification of level of endorsement by OMERACT Community | **○** |
| 19 | Implement communication and dissemination plan | **○** |

# Assembly of working group and protocol development

# 1. Assemble working group

*See Chapter 2 of the OMERACT Handbook for details on assembling an OMERACT Working Group following the Spirit of OMERACT (e.g. collaboration, consensus).*

Place a checkmark here to acknowledge that working group composition meets OMERACT requirements: [ ]

Place a checkmark here to acknowledge that the steering group patient research partners (PRP) will be offered authorship on publications arising from this work: [ ]

(*note that other PRP active in the group may also be offered authorship but at a minimum it is expected that the PRP on the steering group will be offered authorship*)

***Ensure your OMERACT Working Group membership list is up to date on the OMERACT website***

# Decide on methods protocol for Core Outcome Instrument Set selection

## 2.1 Target PICOC of Instrument

*Define in detail the PICOC to which the instrument will apply* [***you will take this from your Core Domain workbook***]

**Domain Definition Report**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Working Group:** | | | **Date completed:** | |
| **Population:** | **Intervention(s):** | **Control(s):  placebo/drug** | | **Context: (target type of study)** |
| **What is the name that you give to your target domain?** |  | **Is this part of a broader domain?** | **|\_\_| No**  **|\_\_| Yes ….If yes, which one \_\_\_\_\_\_\_\_\_\_\_** | |
| **Tell us more about that domain** –this is your domain definition. what is the breadth, depth – what do you want to be able to see. |  | | | |
| Which of the **core areas** does this fall into? (check one) | |\_\_| Pathophysiological manifestations |\_\_| Life impact  |\_\_| Adverse effects |\_\_| Resource use (i.e., costs) | | | |
|  | | | | |
| **Tracking for future reference…. (we suggest you track these now while you are thinking about it, but they are not mandatory. You (or your successors) will need them later in instrument selection)** | | | | |
| ***…How did you come to understand this target domain well?*** | *Qualitative findings (i.e., add relevant quotes from patients, stakeholders that aid in understanding of what this is….and what it is not)* | | | |
| *Other available definitions, frameworks used (i.e., did you take definition from another framework, or another working group – both are great if they work for you – cite here)* | | | |
| ***….Are there any “it all depends” type factors. Factors that make a difference in the number/score obtained?*** | *Examples: +/- use of assistive device, type of imaging machine, technician variability, time of day…* | | | |

## 2.2 Methods to identify candidate instruments

*Describe the methods that will be used for selecting candidate instruments. E.g. scoping or systematic review of RCTs, searching repositories of instruments such as MAPI, COSMIN, grey literature.*

2.3 Description of candidate instrument ***–*** *include details such as theoretical underpinnings, brief summary of content (number of elements, items), response categories, scoring suggestions, how it was developed, what it is intended to measure (as described by the developers of the index/technique/scale)*

***Provide a copy of the instrument below or attached to this workbook. Please specify versions or subscales, if applicable.***

***[INSERT HERE]***

## 2.4 Protocol methods plan

*OMERACT has established an approach to instrument selection. This is the ‘OMERACT Way’ and includes an approach to see if an instrument has passed the OMERACT Filter. If you want to follow our protocol, you can agree to that here and move forward. We encourage you to follow this protocol.*

*Our preference is that working groups follow the approach that has been developed and approved by the OMERACT executive and community. If you are going to deviate from these methods, we need to know ahead of time and the full Technical Advisory Group (TAG) will review.*

**Please mark one of the following two options with an ‘X’:**

[ ] Our working group has reviewed the methods described in this workbook and will follow the OMERACT Instrument Selection Workbook approach including:

• Check for domain match and feasibility using methods described

• Seek agreement of the working group that this is an instrument that matches the target domain & is feasible

• Conduct a literature search using search terms available in the Appendix and modified for your need

• Select articles using screening and selection questions provided and create a PRISMA diagram

• Extract location of the evidence in a Summary of Measurement Properties (SOMP) table

• Conduct a good methods check (quality appraisal)

• Extract data on description of studies and results in our reporting summary tables templates

• Assess adequacy of the results of the measurement property evaluations

• Synthesize results for each measurement property (creating a profile across measurement properties)

* Complete a Summary of Measurement Properties (SOMP) Table to track the evidence

• Apply the OMERACT Algorithm to determine the recommended level of endorsement

• Present the evidence base to the OMERACT community for ratification

**NOTE: If there is no evidence in the literature for the instrument under assessment, please contact your OMERACT methods support person to liaise with the TAG for help with creating protocols to design the methods studies need to create missing evidence. TAG should review protocols for new methods studies [see #12 on the Instrument Selection process checklist].**

[\_] Our working group will alter parts of the OMERACT Methods for Instrument Selection. Please describe these differences in the space below and submit for review and approval of TAG. You will need to wait for endorsement from the full TAG before proceeding (submit to [admin@omeract.org](mailto:admin@omeract.org)).

# Deliverable: Submit protocol using Instrument Selection Workbook to Technical Advisory Group [TAG]

*Please submit this workbook to TAG at* [admin@omeract.org](mailto:admin@omeract.org)

DATE SUBMITTED: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Review and approval of final protocol by TAG

*TAG will review and provide comments on your protocol; note that this may be an iterative process.*

*Once all TAG comments have been addressed and the final protocol is approved by TAG, you can proceed to conduct your methods work and follow the section:* ***Review of evidence of instrument performance for existing or new instrument*** *of the Instrument Selection Checklist to fill in your results.*

DATE APPROVED: ­­­­­­­­­­­­­­­­­­­\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Review of evidence of instrument performance for existing or new instrument

# Part A: Domain match and Feasibility Assessment

# Obtain Working Group and others assessment of match with the target domain

**Is it a match with target domain? (Truth)**

*To answer this question, there are four tasks to complete:*

*1. Survey of working group members about the content and domain match*

*2. Survey of patients and other key stakeholders about the content and domain match*

*3. Review of raw data for this instrument*

*4. Working group comes to a conclusion about match with target domain and content*

## 5. 1 Survey of working group members about the content and domain match

*Sample survey questions are provided in* [*Appendix A*](#_Appendix_A:_Sample)*. You can use any survey software to obtain this information.*

Please provide a summary of your working group’s input regarding the Domain Match of the selected instrument. Below are samples of the types of questions we need you to address; you can use these or similar questions but please provide a summary of your working group’s responses at this level of detail. This example is geared towards PROs; we offer a table with suggestions of how to ask similar types of questions of imaging outcomes, other biomarkers and composite outcomes in [Appendix A](#_Appendix_A:_Sample).

***SUMMARY OF DOMAIN MATCH (IF NECESSARY, REPLACE & PROVIDE YOUR RESULTS HERE):***

**Instrument: ­­**Click here to enter text. **Date:**Click here to enter a date.

|  |  |  |
| --- | --- | --- |
| **Question** | **Working Group’s summary response** | |
| **Is this instrument (think about items, response, domain capture for PROs; for imaging, think about match with domain components) measuring what YOU want to measure? Are the items relevant to your concept, as experienced by your targeted patients’ experiences and for the intended application? Consider sources of variability you identified, are any of those criteria that were considered in the definition of the domain? For example, is using assistive devices permitted in your concept of independence in ADL functioning? Or do you want to specify a particular time of day when you define your concept of pain intensity – night pain, or morning pain for example?**  **Comments:**  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No | N (%)  N (%)  N (%) |
| **Have all important the elements of the target domain for this population, and intended application been included (consider breadth and depth needed)?**  **Comments:**  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No | N (%)  N (%)  N (%) |
| **Is the instrument free of redundant, unnecessary, or potentially inappropriate or sensitive items?**  **Comments:**  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No | N (%)  N (%)  N (%) |
| **Are the items phrased in a clear and understandable way?**  **Comments:**  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No | N (%)  N (%)  N (%) |
| **Are the items written at a level that will be understood by the target population?**  **Comments:**  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No | N (%)  N (%)  N (%) |
| **Are the instructions for completing items and selecting responses for the items clear?**  Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No | N (%)  N (%)  N (%) |
| **Are the response options clear and appropriate for each item (consider match with the question, ordering of responses)?**  **Comments:**  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No | N (%)  N (%)  N (%) |
| **Is the recall period in the instrument appropriate given the population, domain and intended application, i.e, over the past week, last 24 hours (if applicable)?**  **Comments:**  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No  ☐Not applicable | N (%)  N (%)  N (%) |
| **Is the method of scoring appropriate (consider any weighted responses)?**  **Comments:** Click here to enter text. | ☐Yes  ☐Uncertain  ☐No | N (%)  N (%)  N (%) |
| **Final decision based on working group data**(check one)**:**  **☐Good to go**  **☐some cautions but okay**  **☐ not right for this application** | | |

## 5.2 Survey of patients and other key stakeholders about the content and domain match

*Sample survey questions are provided in* [*Appendix A*](#_Appendix_A:_Sample)*. We set up the survey for respondents to assess both domain match and information needed for the next section on feasibility. We have put them together for the practicality of handing patients and others one form to complete. In order to be able to do this well, the respondents will need to know the target domain and its definition from the detailed definition report (part of the core domain set), as well as the intended application (population, setting). Please make sure you let them know this information.*

**Please use the responses to the part of the survey addressing “domain match” to provide the summary results. You can use any survey software to obtain this information.**

***PLACE YOUR SUMMARY RESULTS HERE:***

|  |
| --- |
| **Working Group’s conclusion on respondents’ data** (check one)**:**  **☐Good to go**  **☐some cautions but okay**  **☐ not right for this application** |

## 5.3 Review of Raw Data

*You now need to look at some data on responses to the scale from either a published paper with existing data or from a pilot project the working group undertakes. We encourage working groups to examine data of their own or from publications to look at the distribution of responses, patterns of missing items, or floor and ceiling effects – all indicators of potential problems of the fit of the item content with the population of interest.*

*We suggest you review responses to the instrument from about 30 persons who are similar to your target population. Frequency distributions for each element/item and for each total score can be examined. Gathering data in a sample similar to your target population allows for examination of ceiling (% of respondents with perfect scores) or floor (% at lowest score) that are indicators of missing content range for the respondents. The distribution of the total score in the target population will give a sense of the degree to which it can be treated as a continuous score using parametric statistics (means, t test) rather than non-parametric statistics (medians, ranks). Missing data can indicate sensitive or misunderstood content. Similar features could be examined for each aspect of a composite outcome measure or clinical score or biomarker.*

*This data can inform your decision about the suitability of the content and coverage of the concept. Use the data to really get a sense of what the scale or index is saying. Groups might choose to use a table like the sample table below to summarize responses and provide more detail on the pattern of responses on a multi-item (reflective) scale. The same could be adapted for key features of a composite index. Below is a sample table for presenting this data.*

***Modify the sample table below to fill in your results****.*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Response: Amount of difficulty doing item | | | |  |
| Item # | Item label | Missing (n) | 0 | 1 | 2 | 3 | Mean (0-3) item score |
|  |  |  | No difficulty | Some difficulty | A lot of difficulty | Not able to do |  |
| 1 | Reach | 1 | 148 | 83 | 17 | 1 | 0.48 |
| 2 | Sit | 0 | 127 | 105 | 16 | 1 | 0.56 |
| 3 | Lift | 1 | 117 | 101 | 28 | 2 | 0.66 |
| 4 | work | 3 | 75 | 105 | 57 | 9 | 1.00 |
| 5 | Carry | 2 | 49 | 121 | 70 | 7 | 1.14 |
| 6 | Pain | 3 | 102 | 111 | 33 | 0 | 0.72 |
| 7 | Dress | 4 | 37 | 94 | 80 | 34 | 1.45 |
| 8 | Transport | 2 | 127 | 93 | 26 | 1 | 0.60 |
| 9 | Walk | 1 | 175 | 63 | 10 | 0 | 0.33 |
| 10 | Run | 5 | 153 | 78 | 13 | 0 | 0.43 |
| 11 | Sports | 2 | 130 | 96 | 21 | 0 | 0.56 |
| Instrument: XXXXX  Mean Score:\_\_\_\_\_\_\_\_\_\_\_\_\_\_ SD\_\_\_\_\_\_\_\_\_\_\_\_\_ Median\_\_\_\_\_\_\_\_\_\_\_  Cronbachs Alphs in this data:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | | | | | |

### Review of Raw Data Form

|  |  |  |
| --- | --- | --- |
| **Feature** | **Criterion** | **Score** |
| Check that your data has a good completion rate | >80% of people answered, less than 20% drop out  OR evidence that the responders were similar to the target sample.  **Working Group Comments:**  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Missing data | a) Amount of missing data/responses  b) Pattern of missing – was there any pattern?  **Working Group Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No |
| Floor and Ceiling | Both less than 15%?  **Working Group Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No |
| Perceived completion time | Is it reasonable for intended study?  **Working Group Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No |
| Normality of distribution in target population | Is it reasonable for intended study?  **Working Group Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No |

Results: Domain Match

Based on the Results above (working group survey, respondent’s survey, data review), complete the Working Group’s assessment of whether the instrument is a match with the target domain.

**Red flag (stop, do not continue):\_\_\_\_\_ Amber (some cautions, but continue): \_\_\_\_ Green (good to go):\_\_\_\_\_**



# Obtain Working Group and others assessment of feasibility



*To answer this question, there are three tasks to complete:*

*1. Survey of working group members about the feasibility*

*2. Survey of patients and other key stakeholders about the feasibility*

*3. Working group comes to a conclusion about the feasibility*

*Readily available information on each instrument should be gathered and considered – instructions, costs, copyright, copy of the questionnaire, etc. Contact with the developers often helps with this as does reviewing the manual.*

*With a copy of the instruments and instructions for administration and scoring in hand, the working group can use the example questions below to evaluate the feasibility in the intended setting of the core set.*

## 6.1 Survey of working group members about feasibility

*Sample survey questions are provided in* [*Appendix A*](#_Appendix_A:_Sample)*. You can use any survey software to obtain this information.*

Please provide a summary of your working group’s input regarding the Feasibility of the selected instrument. Below are samples of the types of questions we need you to address; you can use these or similar questions but please provide a summary of your working group’s responses at this level of detail.

***SUMMARY OF FEASIBILITY (IF NECESSARY, REPLACE & PROVIDE YOUR RESULTS HERE):***

**Instrument:**Click here to enter text. **Date:**Click here to enter a date.

|  |  |  |
| --- | --- | --- |
| **Question** | **Working Group’s Summary Response** | |
| **Is it easy for respondents to understand (considering reading level, instructions, health, and literacy needed)?**  **Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No | N (%)  N (%)  N (%) |
| **Can it be completed within a reasonable amount of time given your study context?**  **Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No | N (%)  N (%)  N (%) |
| **Is the method of administration feasible for your application (i.e., computer-based, paper, equipment needs)?**  **Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No | N (%)  N (%)  N (%) |
| **Are the costs feasible? (consider licensing fees, equipment and training costs).**  **Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No | N (%)  N (%)  N (%) |
| **Are the copyright issues (if any) reasonable and manageable?**  **Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No | N (%)  N (%)  N (%) |
| **Are the equipment, space and training needs feasible for you to carry out?**  **Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No | N (%)  N (%)  N (%) |
| **Is it available in the right language/culture for your intended application?**  **Comments:**  Click here to enter text. | ☐ Yes  ☐Uncertain  ☐ No | N (%)  N (%)  N (%) |
| **Final decision based on working group data** (check one)**:**  **☐Good to go**  **☐some cautions but okay**  **☐not right for this application** | | |

## 6.2 Survey of patients and other key stakeholders about feasibility

*It is essential for the working groups to get feedback from the people who will be responding to the instrument (i.e. patients, caregivers, clinicians for a PRO). We suggest you survey 5-10 people to provide good insight into this appraisal of feasibility from the respondent’s perspective. Respondent input (from outside the working group) should be sought to ensure that people like those who will be participating in a study will be giving their opinion.*

*The survey in* [*Appendix A*](#_Appendix_A:_Sample) *was designed to gather information on* ***both*** *the match with domain and the feasibility of this instrument. At this point your group will consider the responses to the questions about feasibility.*

**Please use the responses to the part of the survey addressing “feasibility” to provide the summary results. You can use any survey software to obtain this information.**

***PLACE YOUR SUMMARY RESULTS HERE:***

|  |
| --- |
| **Working Group’s conclusion on respondents’ data** (check one)**:**  **☐Good to go**  **☐some cautions but okay**  **☐ not right for this application** |

Results: Feasibility

Based on the Results above (working group survey, respondent’s survey), complete the Working Group’s assessment of whether the instrument is feasible to use.

**Red flag (stop, do not continue):\_\_\_\_\_ Amber (some cautions, but continue): \_\_\_\_ Green (good to go):\_\_\_\_\_**



# Obtain working group decision based on synthesis of overall ratings of domain match and feasibility

***Working Group’s vote:*** *you should now do a vote across your own working group members and record how you feel about this instrument based on the assessments of domain match and feasibility. This vote is important, and the result should be recorded in this workbook. Groups should achieve at least 70% agreement that this instrument can move forward (that is either a GREEN or AMBER vote). If less than that, the instrument should be set aside.*

*It is common, and a very good practice, to put instruments that are not doing well aside at this point. There is no way to repair or retest a mismatch with the target domain, or a lack of feasibility in using the tool. These instruments should not continue and as shown in the OMERACT Instrument Selection Algorithm they will land in the “STOP” area and not be considered further. This will save you a lot of time, so think this through carefully.*

**Result of Working Group Vote:**

|  |  |  |
| --- | --- | --- |
| Date: | Agree it is a domain match & feasible (%) | Do not agree it is a domain match & feasible (%) |
| Working Group (N= \_\_) |  |  |

Based on the Results above, complete the Working Group’s final assessment of whether the instrument matches the target domain and is feasible to use.

**Green (good to go):\_\_\_\_\_ Amber (some cautions, but continue): \_\_\_\_ Red flag (stop, do not continue):\_\_\_\_\_**

*Complete step 7 on the Instrument Selection checklist. If the decision is to continue to assess the measurement properties of the instrument selection, move on the next section in the workbook, Part B. If the decision is to put the instrument aside, stop here in the workbook. Submit the workbook to* [*admin@omeract.org*](mailto:admin@omeract.org) *to be kept on file.*

|  |  |  |
| --- | --- | --- |
| 7 | **Obtain working group decision based on synthesis of overall ratings of domain match and feasibility:** Is the instrument a match with the domain AND feasible? [*Select Yes or No*]  Yes \_\_\_\_  if yes, continue with Part B of checklist below  No \_\_\_\_  If no, set instrument aside (find new one or develop new one) | **○** |

.

# PART B: Review of the evidence on measurement properties

*This portion of the workbook has been designed to help you, the working groups of OMERACT, move through the review and synthesis of the evidence on measurement properties.*

*One of the most important summaries of your work is a fillable “OMERACT Summary of Measurement Properties” (SOMP) table that is designed to help you to track the literature/evidence you have gathered, how you have appraised it and how your group has synthesized this into a statement of the performance of each instrument. Chapter 5, section 8 of the OMERACT Handbook describes the creation of the SOMP.*

# 8. Conduct literature search to find studies; create PRISMA diagram, fill in Summary of Measurement Properties (SOMP) Table with results of which measurement properties are assessed in each included study

*The first task at this stage is to set up a good literature search to capture the studies on measurement properties for the instrument. At OMERACT, our main focus is on seeing if we have enough evidence to suggest that this Outcome Measurement Instrument can serve as a measure of your target domain in a clinical trial or longitudinal study. We therefore focus on the evidence that will answer whether it meets the OMERACT Filter of TRUTH (construct validity, inter-method reliability) and DISCRIMINATION (test-retest reliability, longitudinal construct validity, discrimination of change in two groups, and thresholds of meaning).*

*Measurement property searches tend to get you a lot of articles that are not about measurement properties. We have provided sample search strategies in* [*Appendix B*](#_Appendix_B:_Sample) *and ways to screen and select articles to try to help with this process.*

*PROSPERO is an international database of prospectively registered systematic reviews. Key features from the review protocol are recorded and maintained as a permanent record. ‘PROSPERO aims to provide a comprehensive listing of systematic reviews registered at inception to help avoid duplication and reduce opportunity for reporting bias by enabling comparison completed review with what was planned in the protocol.’*

*Once finalized, we recommend registering your systematic review protocol in PROSPERO at:* [*https://www.crd.york.ac.uk/PROSPERO/*](https://www.crd.york.ac.uk/PROSPERO/)

Date registered in PROSPERO: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
PROSPERO Registration number: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

8.1. Design a search strategy for the literature search

*The form in* [*Appendix B*](#_Appendix_B:_Sample) *will help you identify the key words to design your search strategy.*

*With the help of an information specialist, choose your databases to identify the peer-reviewed studies in relevant disciplinary areas: MEDLINE and two others are required. Where appropriate, the following databases should be searched: Cumulative Index to Nursing & Allied Health Literature (CINAHL) and PsychInfo. Remember you might have different search terms for each of the databases. The searches should be checked by an information specialist for fit with each database, such that the controlled vocabularies will be utilized when possible in combination with keywords.*

[*Appendix C*](#_Appendix_C._) *provides example search strategies to be used for identifying measurement property studies.*

8.2 Documenting your actual search strategy

*This table should be used to document the specifics of the actual search strategy used. It should be updated with each revision so that a final document demonstrates the full scope of literature that was searched. The date in which the search was conducted, and the dates used in the database, should also be documented. Please enlist the help of an information specialist/librarian to test the search, be sensitive to inclusion of articles at this stage so that important information is not missed.*

*Table 8.1 is an example table for reporting your search****.* *Delete the italic text in the table and fill in with your information.***

**Table 8.1 Search strategy documentation**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date search run:** | | **Dates included in the run (in databases):** | |
| **Database searched:**  *MEDLINE* | | | |
| **Component of PIM** | **Description of your criteria for each component** | | **Search terms used to capture this. (\*May vary by database)** |
| *Population* | *Describe your patient targets as specifically as you need for relevance to your working group.*  *• Clearly define the rheumatological disease(s)/condition(s) of interest.*  *• A broader population can be considered if insufficient evidence is available in target disease/condition.*  *• Any restrictions re age, severity, disease duration should be described.*  *Describe both inclusion and exclusion criteria.*  *Make use of Cochrane Musculoskeletal group for search terms for the diseases.* | | *For osteoarthritis (Wajon et al, 2015):*  *1. exp osteoarthritis/ 2. osteoarthr$.tw. 3. (degenerative adj2 arthritis).tw. 4. arthrosis.tw. 5. or/1-4* |
| *Instrument* | *Use full name as title/abstract, as well as many possible acronyms.*  *• Original articles about the instrument, development, conceptual framework.*  *• Articles using name of instrument or acronym in title/abstract.*  *• Think of all acronyms, what it has been called or might be called.*  *Also conduct a citation search on key articles for instrument, then select to match the other two criteria.*  *e.g. RAID, “Rheumatoid* Arthritis *Impact of Disease”.* | | *“RAID” OR*  *“Rheumatoid Arthritis Impact of Disease”* |
| *Measurement property* | *Use OMERACT-modified Terwee 2009 search engine for properties needed for the Filter 2.2 review (construct validity, test-retest reliability, inter-method reliability, longitudinal construct validity, thresholds of meaning).*  *We are interested in primary articles assessing measurement properties*  *Reviews can be used to identify source articles and review is done only on the source article (primary data).*  *Separate search will also be needed for clinical trials or cohorts using instrument for the property of between-group discrimination in a clinical trial setting. Do this separately from measurement property search for easier management.* | | *Search Terms provided in Appendix A for measurement properties.*  *Use these terms in the database rather than pasting here, you can refer to the version you will use from below list (see appendix).* |

‘**Must have’ articles**

*As a preliminary check of the search strategy’s accuracy, the working group will select what they know to be key articles on measurement properties for the target instrument from personal libraries, developers, instrument websites, or preliminary searches at this phase. They will see if the search strategy above has captured these articles and adjust search strategies to do so. Document changes to your search in the table above. Retain this in accurate form for publication purposes and for future reference by the group. The Working Group can describe if the articles were captured in the table below.*

*Must have articles.* Here are five key articles about this instrument that we think should be captured in a good search strategy.

|  |  |  |
| --- | --- | --- |
|  | Articles that are important to capture in our search. | Was it captured |
| *Eg.* | *Jensen, 1998* | *Yes/No* |
| 1 | Click here to enter text. | ☐Yes ☐No |
| 2 | Click here to enter text. | ☐Yes ☐No |
| 3 | Click here to enter text. | ☐Yes ☐No |
| 4 | Click here to enter text. | ☐Yes ☐No |
| 5 | Click here to enter text.  Click here to enter text. | ☐Yes ☐No |

### 8.3 Screening and selection

#### 8.3.1 Title and Abstract Screening

*Review titles quickly to identify articles that are not about measurement and which can be quickly excluded (responsiveness to a drug, rather than a study of responsiveness for example).*

*Set up in Excel, or software designed for screening (e.g. Distiller or Covidence). Be quick, most will be ruled out.*

*This can be done by a single reviewer if the panel of reviewers has been trained perhaps by screening 15-20 articles together. If groups are concerned by consistency and accuracy, two screeners can be used and inconsistency in responses identified and dealt with by consensus of the two reviewers. Groups should document this in their methods. Below is an example form for screening.*

Is this title/abstract describing a study about one or more measurement properties? (a study about construct validity, test-retest reliability, inter-method reliability, longitudinal construct validity, thresholds of meaning) that includes the XXX instrument?

☐ Yes it is.

☐ Uncertain/unclear. I am not certain based on title/abstract, needs a more thorough review.

☐ No this is not relevant for our review (not measurement, or it is a letter to the editor, abstract for meeting).

☐ Article is in language I cannot read (return to core team for reassignment, or exclusion if outside scope).

#### 8.3.2 Full article review (two reviewers, consensus sought)

*The next stage is a full article review that includes all YES and Uncertain responses at the screening level. Everything with a NO answer should have been excluded.*

*The Full article is reviewed first to confirm that they meet the criteria again (remembering some were checked as UNCLEAR and still need review). It is also the time that we document the measurement properties that were covered in that study.*

*The same three initial questions are asked again, but this time removing the option of saying UNCLEAR as this is as clear as it gets!*

*Examples to help you set up your screening form:*

|  |  |  |
| --- | --- | --- |
| Selection Phase Questions | Responses | Comments |
| Is this a study about a measurement property of the target instrument (as per screening) | ☐ Yes  ☐ No | Click here to enter text. |
| Does this study contain primary data? | ☐ Yes  ☐ No article is a systematic review of measurement properties (will be flagged for extraction of primary articles)  ☐ No, article is not relevant and should not proceed | Click here to enter text. |
| Is it in a rheumatologic patient group or a group you deem (with explanation in your final report) to be close enough to rheumatologic patients that it should be considered? | ☐ Yes  ☐ No (article is NOT relevant to the rheumatological or related population and should NOT be included in this review) | Click here to enter text. |
| Which measurement properties are addressed in this study (ONLY for those studies screening YES to all of above items)? | ☐ Domain match  ☐ Content validity  ☐ Feasibility  ☐ Construct validity/known groups  ☐ Inter-method reliability  ☐ Test retest reliability  ☐ Longitudinal Construct validity  ☐ Discrimination in clinical trial setting  ☐ Thresholds of meaning | Click here to enter text. |

#### 8.3.3 Identifying additional articles

*All articles identified in the screening and all existing systematic reviews should have a hand review of the reference list(s) to identify any additional relevant articles. Repeat with the newly identified relevant articles until saturation. Identify these additional articles as having been found through reference lists of screened articles and reviews. These should be tracked and reported separately from those gathered in the electronic searches in the final report and PRISMA diagram.*

### 8.4 Creating your PRISMA Flow diagram

*When you have completed screening your search, please fill in the blank PRISMA below.*

**PRISMA 2009 Flow Diagram: Measurement Property Reviews**

*Adapted From:*  Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097 **For more information, visit** [**www.prisma-statement.org**](http://www.consort-statement.org/)**.**

New studies conducted to fill gaps (n = )

Full-text articles excluded, with reasons(n = )

Full-text articles assessed for eligibility (n = )

Papers excluded  
(n = )

Screening of title and abstracts (n = )

Records after duplicates removed  
(n = )

Studies covering each measurement property:

Construct validity n=

Inter-method reliability n=   
Test retest reliability n=

Longitudinal construct validity n=

Discrimination in clinical trials n =

Thresholds of meaning n=

**Included**

**Eligibility**

**Screening**

**Identification**

Studies included in qualitative synthesis  
(n = )

Additional records identified through other sources  
(n = )

Records identified through database searching  
(n = )

### 8.5 Creating your Summary of Measurement Property (SOMP) Table

*We have designed a table that will track the findings of the literature you review and will be a key document for presenting your results to OMERACT for review. Here is a summary of how you will be using it over the course of your review.*

*The table is designed to help you track the evidence found in your final set of included articles. Have one row for each article you found in your literature search. Below we have described the overall process of filling in the SOMP Table.*

* *At this stage you have pulled the measurement properties that are in each article. Place an X to indicate which measurement properties are covered in each article. This will help to show you which articles will be your source of information:*
  + *total up number of X in each column to show you the literature to be looked at for each property*
  + *note: a study done by the working group can be added to this table once it is reviewed by TAG. It will be marked as unpublished if not yet published.*
* *In later stages of the review you will update this table to track your progress.*
* *We will tell you about a Good Methods Check (stage of quality appraisal) and at that stage the cells with an X can be filled with the results of that check (RED – serious flaws, risk of bias, do not use; AMBER – some concerns but go ahead with this, and GREEN – good methods used).* 
  + *At that point you can redo your count excluding the RED spaces and this will give you the number of pieces of good quality evidence that are available for each measurement property.*
* *When your data extraction on the results of the assessment of the measurement property is complete, you will change the X to a symbol that summarizes the results of that test:*

*+ = positive support for that measurement property.*

*+/- = ambivalent support, inconclusive result.*

*- = support that this instrument did not reach adequate performance standards for that property.*

* *As you do your measurement property level synthesis, fill the cell in with a red, amber or green colour to track your findings in the bottom row, “Rating (RAGW)”.*
* *Record your group’s proposed overall recommendation for the instrument (Endorsed, Provisional endorsement, Not endorsed) based on the findings in this table in the last row of the table.*

**OMERACT Summary of Measurement Properties Evidence table (a).**    
Place an X to indicate which measurement properties are covered in each included study

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Instrument:  Domain:** | | | | | | **Date completed:** | | | | |
| **Population:** | **Intervention(s):** | | | | **Control:  standard/drug** | | | **Context:  clinical trials** | | |
| **Study  (Author year)** | **Truth**  **Domain match\*** | **Feas-ibility\*** | **Truth** | | | **Discrimination** | | | | |
| Construct validity | Inter-method reliability | | Test retest reliability | Long’l construct validity | | Clinical trial discrimination | Thresholds of meaning |
| **Working Group Appraisal (n= )** |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
|  |  |  |  |  | |  |  | |  |  |
| **Total available studies for each property** |  |  |  |  | |  |  | |  |  |
| **Total studies available for synthesis** |  |  |  |  | |  |  | |  |  |
| **Synthesis Rating** |  |  |  |  | |  |  | |  |  |
| **OMERACT Endorsement** | **Based on the OMERACT Algorithm this instrument is:** | | | | | | | | | |

\*both domain match and feasibility have already been completed by the working group. Any instrument making it this far has already passed these steps. In the SOMP we are tracking if there is any literature published on these topics. It is acceptable that there is no literature on these two items, and we will move forward with the working group’s decision on domain match and feasibility.

# 9. Check to see if each of the included studies has used good methods when assessing each measurement property using the COSMIN-OMERACT good methods check; add these findings into the SOMP Table by coloring cells Green, Amber, or Red.

### 9.1 COSMIN-OMERACT Good Methods Check

*Once you have your articles and their measurement properties organized, you then need to do a “COSMIN-OMERACT Good Methods Check” (i.e. a quality appraisal) on the methods used to evaluate each measurement property in each article.* *Good Methods should be checked by two raters and agreement reached. After the checks have been done, an overall rating is given by the pair of raters to say whether they feel this piece of evidence should go forward for further assessment of the adequacy of the results.*

*Below are the COSMIN-OMERACT Good Methods Checklists for each of the measurement properties in the OMERACT Filter 2.2. Use one table per study; e.g. if you found 3 studies assessing construct validity, you will need 3 of the tables below. In order to help you track your Good Methods Checklist results, we have created a spreadsheet* [(LINK TO EXCEL WORKSHEET)](https://omeract.org/instrument-selection/downloadable-forms/) *based on work by Alessandro Chiarotto who kindly shared his template for us to adapt based on the following reference: Chiarotto A, et al. Measurement properties of Numeric Rating Scale, Visual Analogue Scale and Pain Severity subscale of the Brief Pain Inventory in patients with low back pain, a systematic review. J Pain. 2019 Mar;20(3):245-263.*

*You can use either the Word tables below or the Excel spreadsheet to report the Good Methods Check results.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Pillar: TRUTH**  **Question: Do the numeric scores make sense?**  **Measurement property: Construct (hypothesis testing) validity (COSMIN Space 8)** | | | |
| Author Year | Yes, good methods used | No, not achieved | Notes |
| Was a clear description given of the construct measured by the comparator instrument? | ☐ | ☐ | Click here to enter text. |
| Were the measurement properties of the comparator instrument(s) described and at least adequate? | ☐ | ☐ | Click here to enter text. |
| Were design and statistical methods adequate for the hypotheses to be tested? | ☐ | ☐ | Click here to enter text. |
| Otherwise good methods? (Free of any other important flaws). | ☐ | ☐ | Click here to enter text. |
| Considering the information available, would you recommend this study as evidence to be considered for this measurement property*? (enter this in the OMERACT Summary of Measurement Properties Table)*  **☐ Yes, likely low risk of bias.**  **☐ Some cautions, but this will be used as evidence**  **☐ No, don't use this evidence** | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Pillar: TRUTH**  **Question: Do the numeric scores make sense?**  **Measurement property: Inter-method reliability (e.g. inter-rater, inter-machine)** | | | |
| Author Year | Yes, good methods used | No, not achieved | Notes |
| Were the measurements conducted independently? | ☐ | ☐ | Click here to enter text. |
| Did the design of the study hold all other factors constant except for the source of variability being examined? | ☐ | ☐ | Click here to enter text. |
| Were the test conditions similar for the measurements? (e.g., type of administration, environment, instructions) | ☐ | ☐ | Click here to enter text. |
| Was the correct statistic used?   * Continuous data: intra-class correlation coefficient   (ICC) used.   * Dichotomous/ordinal/nominal scores: Kappa (w) used. | ☐ | ☐ | Click here to enter text. |
| Otherwise good methods? (Free of any other important flaws). | ☐ | ☐ | Click here to enter text. |
| Considering the information available, would you recommend this study as evidence to be considered for this measurement property? *(enter this in the OMERACT Summary of Measurement Properties Table)*  **☐ Yes, likely low risk of bias.**  **☐ Some cautions, but this will be used as evidence**  **☐ No, don't use this evidence** | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Pillar: DISCRIMINATION**  **Question: Can it discriminate between situations of interest?**  **Measurement property: Test-retest reliability (COSMIN Space 5)** | | | |
| Author Year | Yes, good methods used | No, not achieved | Notes |
| Were the patients stable in the interim time period? | ☐ | ☐ | Click here to enter text. |
| Was the time interval appropriate? | ☐ | ☐ | Click here to enter text. |
| Were the test conditions similar for the measurements? (e.g., type of administration, environment, instructions) | ☐ | ☐ | Click here to enter text. |
| Was the correct statistic used?   * Continuous data: intra-class correlation coefficient   (ICC) used.   * Dichotomous/ordinal/nominal scores: Kappa used. | ☐ | ☐ | Click here to enter text. |
| Otherwise good methods? (Free of any other important flaws). | ☐ | ☐ | Click here to enter text. |
| Considering the information available, would you recommend this study as evidence to be considered for this measurement property? *(enter this in the OMERACT Summary of Measurement Properties Table)*  **☐ Yes, likely low risk of bias.**  **☐ Some cautions, but this will be used as evidence**  **☐ No, don't use this evidence** | | | |

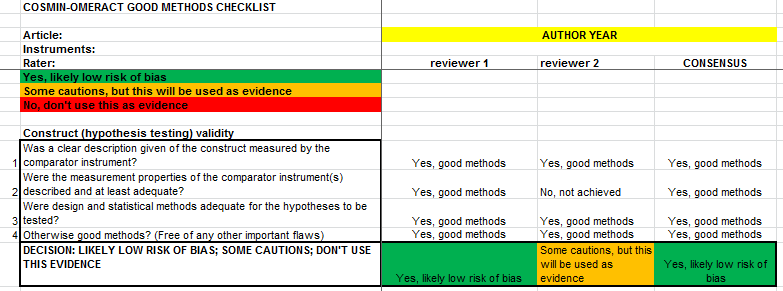
|  |  |  |  |
| --- | --- | --- | --- |
| **Pillar: DISCRIMINATION**  **Question: Can it discriminate between situations of interest?**  **Measurement property: Responsiveness (Longitudinal Construct validity) (COSMIN Space 9 a,b,d)** | | | |
| Author Year | Yes, good methods used | No, not achieved | Notes |
| Can the criterion for change be considered an adequate gold standard OR is the construct for change clear (either as a situation of change or an actual indicator of change)? | ☐ | ☐ | Click here to enter text. |
| Were the measurement properties of the comparator standard described and at least adequate? (N/A for “gold standards). | ☐ | ☐ | Click here to enter text. |
| Were the statistical methods appropriate for the testing situations? (for comparison to gold standard this would include ROC, AUC, predictive values, sensitivity & specificity; correlation of change with external anchor, for constructs: effect size, standardized response mean, correlation). | ☐ | ☐ | Click here to enter text. |
| Otherwise good methods? (Free of any other important flaws). | ☐ | ☐ | Click here to enter text. |
| Considering the information available, would you recommend this study as evidence to be considered for this measurement property? (*enter this in the OMERACT Summary of Measurement Properties Table*)  **☐ Yes, likely low risk of bias.**  **☐ Some cautions, but this will be used as evidence**  **☐ No, don't use this evidence** | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Pillar: DISCRIMINATION**  **Question: Can it discriminate between situations of interest?**  **Measurement property: Clinical trial discrimination (COSMIN Space 9c)** | | | |
| Author Year | Yes, good methods used | No, not achieved | Notes |
| Was the time interval between testing stated and appropriate? | ☐ | ☐ | Click here to enter text. |
| Were there a proportion of people expected to change in one or both groups? (Improvement or deterioration)? | ☐ | ☐ | Click here to enter text. |
| Were hypotheses formulated regarding the anticipated mean differences in change scores between subgroups a priori?   * i.e. positive/negative or no change can be expected. | ☐ | ☐ | Click here to enter text. |
| Were the statistical methods adequate for the hypotheses tested (relative efficiencies, pooled treatment effect sizes, standardized mean differences)? | ☐ | ☐ | Click here to enter text. |
| Otherwise good methods? (Free of any other important flaws). | ☐ | ☐ | Click here to enter text. |
| Considering the information available, would you recommend this study as evidence to be considered for this measurement property? (*enter this in the OMERACT Summary of Measurement Properties Table*)  **☐ Yes, likely low risk of bias.**  **☐ Some cautions, but this will be used as evidence**  **☐ No, don't use this evidence** | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Pillar: DISCRIMINATION**  **Question: Can it discriminate between situations of interest?**  **Measurement property: Thresholds of meaning** | | | |
| Author Year | Yes, good methods used | No, not achieved | Notes |
| Was the patient group similar to your target population (level of disease severity, demographics)? | ☐ | ☐ | Click here to enter text. |
| Is the anchor easily understandable? |  |  |  |
| Is the anchor clearly related to the target domain of interest (i.e. good correlation between anchor and instrument)? |  |  |  |
| Was the cut-off on the anchor used to MID justified to be a small but important difference/important state? |  |  |  |
| Did the same respondent respond to instrument and anchor? |  |  |  |
| Was analysis done separately for improvement and deterioration OR only in same direction anticipated in the target application? | ☐ | ☐ | Click here to enter text. |
| Were multiple criteria and/or analyses used and results triangulated? | ☐ | ☐ | Click here to enter text. |
| Did the analysis include either a Youden index threshold from ROC, or another cut off on an ROC approach? Or if a threshold type of approach (25% or 75%) was used, was it tested for diagnostic utility (sensitivity and specificity)? | ☐ | ☐ | Click here to enter text. |
| Otherwise, good methods? (Free of any other important flaws). | ☐ | ☐ | Click here to enter text. |
| Considering the information available, would you recommend this study as evidence to be considered for this measurement property? (*enter this in the OMERACT Summary of Measurement Properties Table*)  **☐ Yes, likely low risk of bias.**  **☐ Some cautions, but this will be used as evidence**  **☐ No, don't use this evidence** | | | |

*In this spreadsheet you can use colour to track the responses of each rater to the Good Methods Checklist items.*

*Sequential columns show other articles included in this review (same as the rows on the OMERACT Summary of Measurement Properties Evidence Table). An example of one measurement property is listed below.*



### 9.2 Fill in SOMP with results of Good Methods Check

*Colour the cells in the SOMP with the result of each assessment of the Good Methods Check, either GREEN, AMBER, or RED.*

# 10. Conduct data extraction on those measurement properties that were assessed as green or amber Good Methods and complete measurement property tables for summary descriptions of the studies; fill in SOMP Table with assessment of the adequacy of results

### 10.1 Conduct data extraction and summary tables for each measurement property

*The Working Group now needs to create tables to summarize the study characteristics and the actual findings of your evidence. Only measurement property assessments that have passed the COSMIN-OMERACT Good Methods Checklist with AMBER or GREEN ratings will be included. TAG has drafted summary tables to report each measurement property. We have also drafted a table where a general description of each included study can be reported. Groups may choose to format their own tables, but we ask that all the elements in the tables below are included. This includes the study design elements, as well as the analytic approach and results.*

For current versions of summary tables for each of the measurement property tables, please click here:[**https://omeract.org/instrument-selection/downloadable-forms/**](https://omeract.org/instrument-selection/downloadable-forms/)

### 10.2 Judging the PERFORMANCE of the instrument based on the results found in the studies.

*Below are the OMERACT provisional standards for adequate performance. Use this to guide your decisions to complete the judgement of the adequacy section in the summary tables. We use the following symbols:*

*+ = positive support for that measurement property.*

*+/- = ambivalent support, inconclusive result.*

*- = support that this instrument did not reach performance standards for that property.*

*When your data extraction on the results of the assessment of the performance of the measurement property is complete, you will fill in the results in your SOMP Table. Change the X to a symbol that summarizes the results using the* *(+, +/-, -) symbols.*

Copy the completed summary tables into this section of the workbook or submit the completed spreadsheet if you choose to use Excel to record your data.

|  |  |  |
| --- | --- | --- |
| **Pillar (and Question)** | **Measurement property** | **OMERACT Filter 2.2**  **Provisional standards for adequate performance** |
| Truth.  (***Question 3***. Do the numeric scores make sense?) | Internal consistency | Not part of Filter 2.2, if included should be alpha >0.75, higher if target application is individual clinical decision making (0.90). |
| Construct validity | Pre-specified hypotheses are replicated. Should be shown with similar constructs, dissimilar constructs and known groups in order to show both presence and absence of a relationship as appropriate. |
| Inter-method reliability | Intra-class correlation coefficient (ICC); weighted Kappa coefficient (Kw)  Excellent > 0.90.  Good >0.75 (considered adequate for a Green rating)  Excellent needed for measurement if done for individual clinical decision making. Please also report on SEdiff and minimal detectable difference (MDD)-95, Bland-Altman graph is helpful. |
| Discrimination  (***Question 4***: Can it discriminate between groups of interest?) | Test retest reliability | Intra-class correlation coefficient (ICC); weighted Kappa coefficient (Kw)  Excellent > 0.90.  Good >0.75 (considered adequate for a Green rating)  Excellent needed for measurement if done for individual clinical decision making. Please also report on SEdiff and MDD-95, Bland-Altman graph is helpful. |
| Longitudinal construct validity | Consistency with a priori theory in studies that look at situation similar to the intended application. Anticipated large effect expect SRM >0.80, medium/moderate effect, SRM 0.5-0.79, small effect 0.2-0.5. Findings outside the anticipated range should be considered a negative finding. |
| Sensitivity in clinical trials | Longitudinal data are provided for the groups that have changed and separately for groups that have remained stable or had a different amount of change compared to the first group. SRM is greater in change group than in stable or different change group. This difference is also reported in a relative effectiveness statistic (ESgroup12/ESgroup22) = hypothesized magnitude and direction. If reporting on % exceeding a threshold of meaning, please use an empirical cumulative distribution function for each group and highlight the location of your thresholds of meaning. |
| Thresholds of meaning | There are not “standards” for a calculated threshold. We ask only that reporting and context be as clear as possible for users.  Report threshold value and how it was calculated, error boundaries if possible. Thresholds should be related to the anchors used (i.e., threshold for predicting disease activity), sensitivity and specificity of the cut point. For change thresholds, describe relation of both minimal important difference (MID) and minimal detectable change (MDC) and guide interpretation accordingly. |

# 11. Conduct synthesis across evidence available for each measurement property; fill in rating (Green/Amber/Red/White) on SOMP Table

*At this point, your SOMP table shows the results of the Good Methods check using the Red/Amber/Green colour in cells and the adequacy/performance of the results using the symbols “+”, “+/-“ and “-“. Now, the Working Group needs to synthesize the evidence available for each measurement property (i.e. synthesize the evidence down each column). Fill in the row titled “****Synthesis******Rating”****with the Working Group’s assessment of the evidence for each measurement property.*

***Green*** *indicates synthesis of at least 2 studies with good methods showing positive support (“+”) for the measurement property*

***Amber*** *indicates synthesis of only 1 study showing either positive or ambivalent support; or 2 or more studies showing ambivalent support or an inconclusive result*

***Red*** *indicates synthesis of studies with evidence that the instrument did not reach performance standards*

*White indicates no studies assessing this measurement property; i.e. a gap in the evidence*

## 12. Decide if any gaps exist in evidence of measurement properties. If gaps found, draft protocol for new studies to fill gaps. If no gaps, fill in SOMP Table with proposed level of endorsement

### 12. 1 Addressing gaps in the evidence

*The Handbook describes approaches and theories around studies of the measurement properties we are interested in. If you have a gap in the literature, i.e., an absence of evidence for a certain measurement property, it is very appropriate to consult with a senior methodologist and TAG to design a study to fill any gaps. You will submit your protocol for designing this study to TAG (see Step 13). Remember to keep the “good methods checklist” in mind as a guide to the best methods to use in designing your study and you can easily design a study to ensure you get a green rating. Add new studies into your SOMP table (just say “Unpublished” for the year).*

### 12.2 Synthesis to obtain proposed overall rating of the instrument

*Now the working group should have the evidence it needs to make an informed decision about whether this instrument has passed the OMERACT Filter. See the Handbook chapter 5, section 12.3 for details on the synthesis statement. The OMERACT Algorithm to determine the proposed level of endorsement is described in Table 12.4.*

### 12.3 Complete SOMP table with proposed level of endorsement

*To the same SOMP table used to track “Good Methods Check” results, the overall adequacy of the results for each study, and the synthesis for each measurement property, the working group now adds the final synthesis statement in the ‘OMERACT Endorsement’ row.*

*Below is an example completed SOMP table. Delete this example and replace with your completed SOMP table.*

**EXAMPLE Completed Summary of Measurement Properties (SOMP) Table (fictitious)**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Instrument: ABC  Domain: Physical function** | | | | | | | **Date completed: 2021-02-11** | | | | |
| **Population:  rheumatoid arthritis** | | **Intervention(s): drug** | | | | **Control:  placebo/drug** | | | **Type of studies:  clinical trials** | | |
| **Author/year** | **Truth\***  **Domain match** | | **Feasibility\*** | **Truth** | | | **Discrimination** | | | | |
| Construct validity | Inter-method reliability | | Test retest reliability | Long’l construct validity | | Clinical trial discrimination | Thresholds of meaning |
| **Working Group Appraisal (n=20 including 7 PRPs)** | + | | + |  |  | |  |  | |  |  |
| Tugwell 2005 |  | |  | +/– |  | |  | + | |  |  |
| Shea 2004 |  | |  |  |  | |  | + | |  | + |
| Smith 1999 |  | |  |  |  | |  |  | |  |  |
| Beaton 2015 |  | |  |  |  | |  |  | | + |  |
| De Wit 2018 |  | |  |  |  | |  |  | | + |  |
| Wells 2004 |  | |  | + |  | |  |  | |  |  |
| March 2008 |  | |  |  |  | |  |  | | + | +/– |
| D’Agostino 2011 |  | |  |  |  | |  | +/– | |  | + |
| Bingham 2018 |  | |  | + |  | | +/– |  | |  |  |
| Singh 2010 |  | |  | + |  | |  |  | |  |  |
| Strand 2015 |  | |  | +/– |  | |  |  | |  |  |
| Simon 2011 |  | |  |  |  | |  | + | |  | +/– |
| New data from Conaghan 2021 |  | |  |  |  | | + |  | |  |  |
| **Total available studies for each property** | 0 | | 0 | 5 | N/A | | 3 | 5 | | 3 | 4 |
| **Total studies available for synthesis** | 0 | | 0 | 5 | N/A | | 2 | 4 | | 3 | 4 |
| **Synthesis Rating** | **GREEN**  **From Working group** | | **GREEN**  **From Working group** | **GREEN** | **N/A** | | **AMBER** | **GREEN** | | **GREEN** | **AMBER** |
| **OMERACT Endorsement** | **Based on the OMERACT algorithm this instrument is:**  **Provisionally endorsed**  ***More work needed on test-retest reliability and thresholds of meaning.*** | | | | | | | | | | |

\*both domain match and feasibility have already been completed by the working group. Any instrument making it this far has already passed these steps. In the SOMP we are tracking if there is any literature published on these topics. It is acceptable that there is no literature on these two items, and we will move forward with the working group’s decision on domain match and feasibility.

## 13. Deliverable: Submit the Instrument Selection Workbook to TAG

### 13.1 Assemble documentation for submission

*Now you need to assemble the documentation you have gathered in this workbook and compile a report on the findings and decisions and submit this to the Technical Advisory Group (TAG) at* [*admin@omeract.org*](mailto:admin@omeract.org)

*The report will be the source document for your findings, for future publications and as a reliable foundation for future updates.*

*The completed workbook should include:*

* *Statement of domain definition and intended context of use (the PICOC)*
* *Copy of instrument, and how you found it, brief description of it*
* *Summary of domain match including working group and respondents’ appraisals and the summary check of a sample of raw data.*
* *Summary of decisions and working group vote for domain match and feasibility*
* *Search strategy, dates searched and databases*
* *PRISMA diagram*
* *Excel sheet or equivalent summarizing COSMIN-OMERACT Good Methods Check*
* *Summary description of findings from studies using good methods (templates available)*
* *Summary of Measurement Property Evidence table in final form*

DATE SUBMITTED: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

### 13.2 Draft and submit protocol for new measurement property study

*As described in step 12, if there is a gap in the evidence for a measurement property or properties, draft a protocol describing the study design and, if applicable, a priori hypotheses. Remember to keep the “good methods checklist” in mind as a guide to the best methods to use in designing your study and you can easily design a study to ensure you get a green rating.*

DATE SUBMITTED: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

If no protocol for a new measurement study is being submitted, mark here with an ‘x’: [ ]

## 14. Receive final response from TAG

*TAG members will review the report and ask you for points of clarification. This may be an iterative process until all the TAG questions have been addressed.*

DATE TAG SIGNS OFF ON FINAL RESPONSES FROM WORKING GROUP: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## 15. If studies are needed to fill gaps, conduct new measurement property studies, submit to TAG for Good Methods check, add to body of evidence (SOMP) and go back to Step 12

## If no studies are needed, put X here: \_\_\_\_\_\_and move to Step 16

*TAG will review the results of new measurement property studies conducted by the working group and conduct a ‘Good Methods check’. They will add their assessment to the ‘Good Methods Check’ section in Step 9. Once they have completed this check, the working group will add this study (or studies) to the SOMP and resubmit the workbook as per Step 12.*

IF APPLICABLE, DATE OF RESUBMISSION OF WORKBOOK TO TAG:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## 16. Obtain final agreement on final report

*Submit the final report to TAG including this workbook and all accompanying documentation. The TAG will provide feedback to the working group and feedback will be integrated until the TAG and the working group agree the report is final report.*

DATE TAG AND WORKING GROUP AGREE ON FINAL REPORT:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## 17. Set timeline for next review of instrument

*Describe below your plans/timeline for the next updated review of the evidence for this instrument. Note that OMERACT suggests a maximum of 10 years duration until the next update.*

## 18. Ratification of level of endorsement by OMERACT Community and communication of results

*Consistent with the culture of OMERACT, we bring our work back to a wider community for final endorsement. Working groups will use an online platform supplied by OMERACT and make the following items available to the wider community:*

* *Domain definition worksheet*
* *PRISMA Flowchart*
* *Summary (data extraction) tables*
* *Completed SOMP*

*Working group members will monitor an online discussion board for 2 weeks to allow the community to ask questions about the results leading to the proposed level of endorsement.*

*Online meeting: the working group will host an online meeting (ideally at two separate times to allow for all time zones to be included) in order to summarize the results and address any key questions that arose on the discussion board.*

*Following the feedback session attendees will be sent a voting survey asking them is they agree with the proposed recommendation that, based on the evidence presented, the [named instrument] be given a [Full endorsement, Provisional endorsement, Not endorsed at this time] for the domain of [x] in persons with [Condition]? Yes/No*

*The results of this survey will be the record of endorsement.*

*If over 70% agree, the proposed level of endorsement will become the ratified level of endorsement.*

DATE OF ONLINE MEETING: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

VOTING RESULT: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## 19. Implement communication and dissemination plan

*Those working groups who have completed this instrument selection process will be invited to present their work on a poster at the next OMERACT meeting. Implement your end of project dissemination plan. Outline your key dissemination plans here. We have put some questions from the* [*Tunis et al (JRheum 2017)*](https://www.jrheum.org/content/jrheum/early/2017/07/26/jrheum.161273.full.pdf) *paper on knowledge translation to help you think through the process:*

*• What are the key implications of the research/core set and its adoption (formulating key messages, how and*

*when the results can be used)?*

*• Where will this be disseminated (e.g., journals, events)?*

*• How will the information be communicated to different audiences and end users?*

*• Who are the most effective communicators that you will engage?*

Please outline your dissemination plans:

# Acknowledgements.

This document is built on the experience and a workbook used by the Systematic Review Group at the Institute for Work & Health in Toronto. Special thanks to Ms Emma Irvin, and Drs Kim Cullen and Dr Ben Amick for sharing their workbook and advice. Special thanks to the Measurement Group at the Institute for Work & Health for helping with an in-depth review of reviews of measurement properties that provided a basis for our decisions over quality assessment and adequacy.

This document was prepared by the Handbook Committee – Dorcas Beaton, Peter Tugwell, Beverley Shea, Lara Maxwell, Shawna Grosskleg, Maarten Boers and George Wells. We thank many people, including the executive of OMERACT and Caroline Terwee from COSMIN for thoughtful input in the past – and in the future!

# Appendix A: Sample survey questions for surveying patients/other respondents on domain match and feasibility

The sample questions below are based on assessing domain match and feasibility of a PRO. They can be modified for use with other types of instruments using the considerations in this table:

|  |  |  |  |
| --- | --- | --- | --- |
| **PRO’s** | **Composites** | **Imaging Outcome Instrument (scoring system)** | **Other biomarkers (e.g. ESR, CRP)** |
| Match to domain definition | Is there a clear match with each domain to the target of the composite (i.e., disease activity?) | Does it capture the target domain? | Is this biomarker a good match to the domain definition? |
| Framing of the domain. Do the instructions for the instrument orient the respondent/  observer as to how to consider important sources of variability? | Does the questionnaire specify how to manage things like assistance, assistive devices, or coping strategies when answering? | Does the imaging technique description offer choices to avoid specific sources of variability? i.e., T2 weighting in MRI, or specific angle used for an Xray examination. | Does the instrument offer specific directions that help to avoid variability in scores?  For example, blood pressure can vary by time of day, examiner, and environmental factors. |
| Do items cover the essential elements of the domain from detailed definition template? | Coverage of key elements of the target composite domain in the parts of the composite | Coverage of the elements of that domain | Is it capturing all the essential elements as described in the domain definition? |
| Response options | Scoring of each of the parts of the composite (remember inherent weighting given by the scaling of each domain) | Scoring of each element appropriate? | How is it quantified and is that standardized? |
| Weighting of items into score | How are elements weighted in the composite scoring system? | How the scoring system weights the elemental components in the score? | How is it scored compared to norms? |

**Feedback from respondents to instrument:**

**Is it a match with target domain? (Truth) & Is it practical to use? (Feasibility)**

**Instrument:**Click here to enter text. **Date:**Click here to enter a date.

|  |  |
| --- | --- |
| ***Match to Domain: Thinking about the content of the actual questions/items in the instrument, based on experience of this domain.*** | Respondents answer |
| Are the items in this instrument relevant to you and your experience?  Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Do you think there should be any additional items (i.e., were there things that were missed)? Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Do you think that there should be any items taken out of the instrument?  If yes, tell us why.  Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Were there overlapping, sensitive, or embarrassing items?  Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Does the instrument overall reflect your experience of your [domain]?  Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Did you find that all the items were easy to read? If not, which items were not easy to read?  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Did you feel that all the items were clear and understandable? Could you understand what all the questions were trying to ask? If not, which items did you feel were unclear?  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Did you think that the response options were clear and understandable (i.e. did the possible answers match well with the questions)? If not, which items did you feel had a mismatched response scale?)  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Were the instructions for answering the items clear?  Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Does the timing of the recall period seem reasonable to you (e.g. over the past week, last 24 hours) (if applicable)?  Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No  ☐Not applicable |
| ***Feasibility: Questions about the practical considerations about this instrument.*** | Respondents answer |
| Was it easy enough to complete?  Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Did it take a reasonable amount of time to complete?  Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |
| Did the format seem appropriate (how it looked on the page, font size, how items and responses were organized)? | ☐Yes  ☐Uncertain  ☐No |
| Do you think there was too much equipment and training needed before you could be able to respond to this instrument?  Comments:  Click here to enter text. | ☐Yes  ☐Uncertain  ☐No |

# Appendix B: Sample form to help design your search strategy

|  |  |  |
| --- | --- | --- |
| Use this table to help describe what you want to capture in your literature search. Do not worry about search terms yet, this is your broad description of what you want to look for.  Defining the attributes to be used in the search strategy: Population, Instrument and Measurement properties (PIM). | | |
|  | Guidance from OMERACT FILTER INSTRUMENT SELECTION ALGORITHM | YOUR STUDY (example below in italics, delete and replace with yours; can use this for PROSPERO registration): |
| Population | * Clearly define the rheumatological disease(s)/condition(s) of interest. * A broader population can be considered if insufficient evidence is available in target disease/condition. * Any restrictions re age, severity, disease duration should be described. | * *Low back pain (simple)* |
| Instrument of interest (IOI) | * Original articles about the instrument, development, conceptual framework. * Articles using name of instrument or acronym in title/abstract. * Think of all acronyms, what it has been called or might be called. | *Pain intensity numeric rating scale* |
| Measurement properties of interest | * Use modified Terwee (2009) search engine for properties needed for the Filter 2.1 review (construct validity, test-retest reliability, longitudinal construct validity, clinical trial discrimination, and thresholds of meaning). * Separate search will also be needed for clinical trials or cohorts using instrument for the property of discrimination in a clinical trial setting. * Primary articles are priority. * Reviews can be used to identify source articles and review is done only on the source article (primary data). | * As recommended in the column to the left |
| Publication Language to be included | * Languages should be identified for search. English should be included due to nature of published literature, and any other languages the group members can manage. * Groups must define their scope and how many languages they have the capacity to critique. | * *English, French, Dutch, German* |
| Databases (we recommend using multiple sources for measurement articles. Ideally 3). | * Name the databases that will be searched for the measurement articles; three is recommended to ensure sensitivity in search. * MEDLINE should be included. * Should relate to the domain and instrument of interest (i.e., if it is depression scale, psychological database should be searched). | * *MEDLINE, PubMed, Embase* |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Appendix C. Search Strategies to be used for measurement property studies. **Search terms for the population, instrument and measurement property (PIM) moved from table above into actual search strings. They vary for each target database. Here the example is for low back pain and a pain numeric rating scale. Section 3 about the measurement properties terms should remain constant for all searches and is mirrored off of the Terwee 2009 strategy. The pain examples are included to show things to think about for your own search. Special thanks to the Pain NRS review groups – Alessandro Chiarotto and Lara Maxwell for sharing this search strategy.**  **Search Strategy in MEDLINE**  **#1 POPULATION:**  "Back Pain"[Mesh] OR back pain[tiab] OR lbp[tw] OR "Back Injuries"[Mesh] OR back injury[tiab] OR "Lumbar Vertebrae"[Mesh] OR lumbar spine[tiab] OR "Sacroiliac Joint"[Mesh] OR sacroiliac pain[tiab] OR "Zygapophyseal Joint"[Mesh] OR facet pain[tiab] OR "Sciatica"[Mesh] OR sciatica[tiab] OR "Spinal Stenosis"[Mesh] OR lumbar stenosis[tiab]OR "Intervertebral Disc Displacement"[Mesh] OR herniated dis\*[tiab] OR disc pain[tiab] OR disk pain[tiab] OR "Spondylolisthesis"[Mesh] OR spondylolisthesis[tiab]  **#2: INSTRUMENT:** (numeric\*[tiab] AND rating [tiab] AND (scale[tiab] OR score[tiab])) OR numeric scale[tiab] OR nrs[tw] OR nprs[tw]   **#3 MEASUREMENT PROPERTIES:**   (instrumentation[sh] OR methods[sh] OR "Validation Studies"[pt] OR "Comparative Study"[pt] OR "psychometrics"[MeSH] OR psychometr\*[tiab] OR clinimetr\*[tw] OR clinometr\*[tw] OR "outcome assessment (health care)"[MeSH] OR "outcome assessment"[tiab] OR "outcome measure\*"[tw] OR "observer variation"[MeSH] OR "observer variation"[tiab] OR "Health Status Indicators"[Mesh] OR "reproducibility of results"[MeSH] OR reproducib\*[tiab] OR "discriminant analysis"[MeSH] OR reliab\*[tiab] OR unreliab\*[tiab] OR valid\*[tiab] OR "coefficient of variation"[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR "internal consistency"[tiab] OR (cronbach\*[tiab] AND (alpha[tiab] OR alphas[tiab])) OR (item[tiab] AND (correlation\*[tiab] OR selection\*[tiab] OR reduction\*[tiab])) OR agreement[tw] OR precision[tw] OR imprecision[tw] OR "precise values"[tw] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab\*[tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab\*[tw] OR ((replicab\*[tw] OR repeated[tw]) AND (measure[tw] OR measures[tw] OR findings[tw] OR result[tw] OR results[tw] OR test[tw] OR tests[tw])) OR generaliza\*[tiab] OR generalisa\*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation\*[tiab]) OR discriminative[tiab] OR "known group"[tiab] OR "factor analysis"[tiab] OR "factor analyses"[tiab] OR "factor structure"[tiab] OR "factor structures"[tiab] OR dimension\*[tiab] OR subscale\*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR "item discriminant"[tiab] OR "interscale correlation\*"[tiab] OR error[tiab] OR errors[tiab] OR "individual variability"[tiab] OR "interval variability"[tiab] OR "rate variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv\*[tiab] OR responsive\*[tiab] OR (limit[tiab] AND detection[tiab]) OR "minimal detectable concentration"[tiab] OR interpretab\*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small\*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR "meaningful change"[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab])  #4  ("addresses"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "directory"[Publication Type] OR "editorial"[Publication Type] OR "festschrift"[Publication Type] OR "interview"[Publication Type] OR "lectures"[Publication Type] OR "legal cases"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "patient education handout"[Publication Type] OR "popular works"[Publication Type] OR "congresses"[Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference, nih"[Publication Type] OR "practice guideline"[Publication Type]) NOT ("animals"[MeSH Terms] NOT "humans"[MeSH Terms])  #5 #1 AND #2 AND #3  #6 #5 NOT #4  **Search Strategy in Embase**  #1  (‘numeric’:ab,ti OR ‘numerical’:ab,ti) AND ‘rating’:ab,ti AND (‘scale’:ab,ti OR ‘score’:ab,ti)) OR ‘numeric scale’:ab,ti OR nrs OR nprs  #2 'backache'/exp OR ‘back pain’:ab,ti OR lbp OR 'lumbar spine'/exp OR ‘lumbar spine’:ab,ti OR ‘lumbar spine’:ab,ti OR ‘sacroiliac joint’/exp OR ‘sacroiliac pain’:ab,ti OR ‘zygapophyseal joint’/exp OR ‘zygapophyseal pain’:ab,ti OR ‘facet pain’:ab,ti OR 'sciatica'/exp OR ‘sciatica’:ab,ti OR 'lumbar spinal stenosis'/exp OR ‘lumbar stenosis’:ab,ti OR 'lumbar disk hernia'/exp OR ‘lumbar hernia’:ab,ti OR ‘disk pain’:ab,ti OR 'spondylolisthesis'/exp OR ‘spondylolisthesis’:ab,ti AND [embase]/lim  #3 'intermethod comparison'/exp OR 'data collection method'/exp OR 'validation study'/exp OR 'feasibility study'/exp OR 'pilot study'/exp OR 'psychometry'/exp OR 'reproducibility'/exp OR reproducib\*:ab,ti OR 'audit':ab,ti OR psychometr\*:ab,ti OR clinimetr\*:ab,ti OR clinometr\*:ab,ti OR 'observer variation'/exp OR 'observer variation':ab,ti OR 'discriminant analysis'/exp OR 'validity'/exp OR reliab\*:ab,ti OR valid\*:ab,ti OR 'coefficient':ab,ti OR 'internal consistency':ab,ti OR (cronbach\*:ab,ti AND ('alpha':ab,ti OR 'alphas':ab,ti)) OR 'item correlation':ab,ti OR 'item correlations':ab,ti OR 'item selection':ab,ti OR 'item selections':ab,ti OR 'item reduction':ab,ti OR 'item reductions':ab,ti OR 'agreement':ab,ti OR 'precision':ab,ti OR 'imprecision':ab,ti OR 'precise values':ab,ti OR 'test-retest':ab,ti OR ('test':ab,ti AND 'retest':ab,ti) OR (reliab\*:ab,ti AND ('test':ab,ti OR 'retest':ab,ti)) OR 'stability':ab,ti OR 'interrater':ab,ti OR 'inter-rater':ab,ti OR 'intrarater':ab,ti OR 'intra-rater':ab,ti OR 'intertester':ab,ti OR 'inter-tester':ab,ti OR 'intratester':ab,ti OR 'intra-tester':ab,ti OR 'interobeserver':ab,ti OR 'inter-observer':ab,ti OR 'intraobserver':ab,ti OR 'intra-observer':ab,ti OR 'intertechnician':ab,ti OR 'inter-technician':ab,ti OR 'intratechnician':ab,ti OR 'intra-technician':ab,ti OR 'interexaminer':ab,ti OR 'inter-examiner':ab,ti OR 'intraexaminer':ab,ti OR 'intra-examiner':ab,ti OR 'interassay':ab,ti OR 'inter-assay':ab,ti OR 'intraassay':ab,ti OR 'intra-assay':ab,ti OR 'interindividual':ab,ti OR 'inter-individual':ab,ti OR 'intraindividual':ab,ti OR 'intra-individual':ab,ti OR 'interparticipant':ab,ti OR 'inter-participant':ab,ti OR 'intraparticipant':ab,ti OR 'intra-participant':ab,ti OR 'kappa':ab,ti OR 'kappas':ab,ti OR 'coefficient of variation':ab,ti OR repeatab\*:ab,ti OR (replicab\*:ab,ti OR 'repeated':ab,ti AND ('measure':ab,ti OR 'measures':ab,ti OR 'findings':ab,ti OR 'result':ab,ti OR 'results':ab,ti OR 'test':ab,ti OR 'tests':ab,ti)) OR generaliza\*:ab,ti OR generalisa\*:ab,ti OR 'concordance':ab,ti OR ('intraclass':ab,ti AND correlation\*:ab,ti) OR 'discriminative':ab,ti OR 'known group':ab,ti OR 'factor analysis':ab,ti OR 'factor analyses':ab,ti OR 'factor structure':ab,ti OR 'factor structures':ab,ti OR 'dimensionality':ab,ti OR subscale\*:ab,ti OR 'multitrait scaling analysis':ab,ti OR 'multitrait scaling analyses':ab,ti OR 'item discriminant':ab,ti OR 'interscale correlation':ab,ti OR 'interscale correlations':ab,ti OR ('error':ab,ti OR 'errors':ab,ti AND (measure\*:ab,ti OR correlat\*:ab,ti OR evaluat\*:ab,ti OR 'accuracy':ab,ti OR 'accurate':ab,ti OR 'precision':ab,ti OR 'mean':ab,ti)) OR 'individual variability':ab,ti OR 'interval variability':ab,ti OR 'rate variability':ab,ti OR 'variability analysis':ab,ti OR ('uncertainty':ab,ti AND ('measurement':ab,ti OR 'measuring':ab,ti)) OR 'standard error of measurement':ab,ti OR sensitiv\*:ab,ti OR responsive\*:ab,ti OR ('limit':ab,ti AND 'detection':ab,ti) OR 'minimal detectable concentration':ab,ti OR interpretab\*:ab,ti OR (small\*:ab,ti AND ('real':ab,ti OR 'detectable':ab,ti) AND ('change':ab,ti OR 'difference':ab,ti)) OR 'meaningful change':ab,ti OR 'minimal important change':ab,ti OR 'minimal important difference':ab,ti OR 'minimally important change':ab,ti OR 'minimally important difference':ab,ti OR 'minimal detectable change':ab,ti OR 'minimal detectable difference':ab,ti OR 'minimally detectable change':ab,ti OR 'minimally detectable difference':ab,ti OR 'minimal real change':ab,ti OR 'minimal real difference':ab,ti OR 'minimally real change':ab,ti OR 'minimally real difference':ab,ti OR 'ceiling effect':ab,ti OR 'floor effect':ab,ti OR 'item response model':ab,ti OR 'irt':ab,ti OR 'rasch':ab,ti OR 'differential item functioning':ab,ti OR 'dif':ab,ti OR 'computer adaptive testing':ab,ti OR 'item bank':ab,ti OR 'cross-cultural equivalence':ab,ti AND [embase]/lim  #4 #1 AND #2 AND #3  **Search Strategy in CINAHL**  #1 ((TI numeric\* OR AB numeric\*) AND (TI rating OR AB rating) AND (TI scale OR AB scale OR TI score OR AB score)) OR (TI numeric scale OR AB numeric scale) OR nrs OR nprs  #2 (MH “Back Pain+”) OR TI back pain OR AB back pain OR lbp OR (MH “Back Injuries+”) OR TI back injury OR AB back injury OR (MH “Lumbar Vertebrae”) OR TI lumbar spine OR AB lumbar spine OR (MH “Sacroiliac Joint”) OR (MH “Sacroiliac Joint Dysfunction”) OR TI sacroiliac pain OR AB sacroiliac pain OR (MH “Zygapophyseal Joint”) OR TI facet pain OR AB facet pain OR (MH “Sciatica”) OR TI sciatica OR AB sciatica OR (MH “Spinal Stenosis”) OR TI lumbar stenosis OR AB lumbar stenosis OR (MH “Intervertebral Disk Displacement”) OR TI herniated disk OR AB herniated disk OR TI disk pain OR AB disk pain OR (MH “Spondylosis+”) OR TI spondylolisthesis OR AB spondylolisthesis  #3 TI psychometr\* OR TI observer variation OR TI reproducib\* OR TI reliab\* OR TI unreliab\* OR TI valid\* OR TI coefficient OR TI homogeneity OR TI homogeneous OR TI “internal consistency” OR AB psychometr\* OR AB observer variation OR AB reproducib\* OR AB reliab\* OR AB unreliab\* OR AB valid\* OR AB coefficient OR AB homogeneity OR AB homogeneous OR AB “internal consistency” OR (TI cronbach\* OR AB cronbach\* AND (TI alpha OR AB alpha OR TI alphas OR AB alphas)) OR (TI item OR AB item AND (TI correlation\* OR AB correlation\* OR TI selection\* OR AB selection\* OR TI reduction\* OR AB reduction\*)) OR TI agreement OR TI precision OR TI imprecision OR TI “precise values” OR TI test-retest OR AB agreement OR AB precision OR AB imprecision OR AB “precise values” OR AB test-retest OR (TI test OR AB test AND TI retest OR AB retest) OR (TI reliab\* OR AB reliab\* AND (TI test OR AB test OR TI retest or AB retest)) OR TI stability OR TI interrater OR TI inter-rater OR TI intrarater OR TI intra-rater OR TI intertester OR TI inter-tester OR TI intratester OR TI intra-tester OR TI interobserver OR TI inter-observer OR TI intraobserver OR TI intra-observer OR TI intertechnician OR TI inter-technician OR TI intratechnician OR TI intra-technician OR TI interexaminer OR TI inter-examiner OR TI intraexaminer OR TI intra-examiner OR TI interassay OR TI inter-assay OR TI intraassay OR TI intra-assay OR TI interindividual OR TI inter-individual OR TI intraindividual OR TI intra-individual OR TI interparticipant OR TI inter-participant OR TI intraparticipant OR TI intra-participant OR TI kappa OR TI kappa’s OR TI kappas OR TI repeatab\* OR AB stability OR AB interrater OR AB inter-rater OR AB intrarater OR AB intra-rater OR AB intertester OR AB inter-tester OR AB intratester OR AB intra-tester OR AB interobserver OR AB inter-observer OR AB intraobserver OR AB intra-observer OR AB intertechnician OR AB inter-technician OR AB intratechnician OR AB intra-technician OR AB interexaminer OR AB inter-examiner OR AB intraexaminer OR AB intra-examiner OR AB interassay OR AB inter-assay OR AB intraassay OR AB intra-assay OR AB interindividual OR AB inter-individual OR AB intraindividual OR AB intra-individual OR AB interparticipant OR AB inter-participant OR AB intraparticipant OR AB intra-participant OR AB kappa OR AB kappa’s OR AB kappas OR AB repeatab\* OR ((TI replicab\* OR AB replicab\* OR TI repeated OR AB repeated) AND (TI measure OR AB measure OR TI measures OR AB measures OR TI findings OR AB findings OR TI result OR AB result OR TI results OR AB results OR TI test OR AB test OR TI tests OR AB tests)) OR TI generaliza\* OR TI generalisa\* OR TI concordance OR AB generaliza\* OR AB generalisa\* OR AB concordance OR (TI intraclass OR AB intraclass AND TI correlation\* or AB correlation\*) OR TI discriminative OR TI “known group” OR TI factor analysis OR TI factor analyses OR TI dimension\* OR TI subscale\* OR AB discriminative OR AB “known group” OR AB factor analysis OR AB factor analyses OR AB dimension\* OR AB subscale\* OR (TI multitrait OR AB multitrait AND TI scaling OR AB scaling AND (TI analysis OR AB analysis OR TI analyses OR AB analyses)) OR TI item discriminant OR TI interscale correlation\* OR TI error OR TI errors OR TI “individual variability” OR AB item discriminant OR AB interscale correlation\* OR AB error OR AB errors OR AB “individual variability” OR (TI variability OR AB variability AND (TI analysis OR AB analysis OR TI values OR AB values)) OR (TI uncertainty OR AB uncertainty AND (TI measurement OR AB measurement OR TI measuring OR AB measuring)) OR TI “standard error of measurement” OR TI sensitiv\* OR TI responsive\* OR AB “standard error of measurement” OR AB sensitiv\* OR AB responsive\* OR ((TI minimal OR TI minimally OR TI clinical OR TI clinically OR AB minimal OR AB minimally OR AB clinical OR AB clinically) AND (TI important OR TI significant OR TI detectable OR AB important OR AB significant OR AB detectable) AND (TI change OR AB change OR TI difference OR AB difference)) OR (TI small\* OR AB small\* AND (TI real OR AB real OR TI detectable OR AB detectable) AND (TI change OR AB change OR TI difference OR AB difference)) OR TI meaningful change OR TI “ceiling effect” OR TI “floor effect” OR TI “Item response model” OR TI IRT OR TI Rasch OR TI “Differential item functioning” OR TI DIF OR TI “computer adaptive testing” OR TI “item bank” OR TI “cross-cultural equivalence” OR TI outcome assessment OR AB meaningful change OR AB “ceiling effect” OR AB “floor effect” OR AB “Item response model” OR AB IRT OR AB Rasch OR AB “Differential item  #4 #1 AND #2 AND #3  **Search Strategy in SportDiscus**  #1 ((TI numeric\* OR AB numeric\*) AND (TI rating OR AB rating) AND (TI scale OR AB scale OR TI score OR AB score)) OR (TI numeric scale OR AB numeric scale) OR nrs OR nprs  #2 DE "BACKACHE" OR TI back pain OR AB back pain OR lbp OR TI back injury OR AB back injury OR DE "LUMBAR vertebrae" OR TI lumbar spine OR AB lumbar spine OR DE "SACROILIAC joint" OR TI sacroiliac pain OR AB sacroiliac pain OR DE "ZYGAPOPHYSEAL joint" OR TI facet pain OR AB facet pain OR DE "SCIATICA" OR TI sciatica OR AB sciatica OR (DE "SPINAL canal -- Stenosis") OR TI lumbar stenosis OR AB lumbar stenosis OR DE "INTERVERTEBRAL disk displacement" OR TI herniated disk OR AB herniated disk OR TI disk pain OR AB disk pain OR DE "SPONDYLOLISTHESIS" OR TI spondylolisthesis OR AB spondylolisthesis  #3 TI psychometr\* OR TI observer variation OR TI reproducib\* OR TI reliab\* OR TI unreliab\* OR TI valid\* OR TI coefficient OR TI homogeneity OR TI homogeneous OR TI “internal consistency” OR AB psychometr\* OR AB observer variation OR AB reproducib\* OR AB reliab\* OR AB unreliab\* OR AB valid\* OR AB coefficient OR AB homogeneity OR AB homogeneous OR AB “internal consistency” OR (TI cronbach\* OR AB cronbach\* AND (TI alpha OR AB alpha OR TI alphas OR AB alphas)) OR (TI item OR AB item AND (TI correlation\* OR AB correlation\* OR TI selection\* OR AB selection\* OR TI reduction\* OR AB reduction\*)) OR TI agreement OR TI precision OR TI imprecision OR TI “precise values” OR TI test-retest OR AB agreement OR AB precision OR AB imprecision OR AB “precise values” OR AB test-retest OR (TI test OR AB test AND TI retest OR AB retest) OR (TI reliab\* OR AB reliab\* AND (TI test OR AB test OR TI retest or AB retest)) OR TI stability OR TI interrater OR TI inter-rater OR TI intrarater OR TI intra-rater OR TI intertester OR TI inter-tester OR TI intratester OR TI intra-tester OR TI interobserver OR TI inter-observer OR TI intraobserver OR TI intra-observer OR TI intertechnician OR TI inter-technician OR TI intratechnician OR TI intra-technician OR TI interexaminer OR TI inter-examiner OR TI intraexaminer OR TI intra-examiner OR TI interassay OR TI inter-assay OR TI intraassay OR TI intra-assay OR TI interindividual OR TI inter-individual OR TI intraindividual OR TI intra-individual OR TI interparticipant OR TI inter-participant OR TI intraparticipant OR TI intra-participant OR TI kappa OR TI kappa’s OR TI kappas OR TI repeatab\* OR AB stability OR AB interrater OR AB inter-rater OR AB intrarater OR AB intra-rater OR AB intertester OR AB inter-tester OR AB intratester OR AB intra-tester OR AB interobserver OR AB inter-observer OR AB intraobserver OR AB intra-observer OR AB intertechnician OR AB inter-technician OR AB intratechnician OR AB intra-technician OR AB interexaminer OR AB inter-examiner OR AB intraexaminer OR AB intra-examiner OR AB interassay OR AB inter-assay OR AB intraassay OR AB intra-assay OR AB interindividual OR AB inter-individual OR AB intraindividual OR AB intra-individual OR AB interparticipant OR AB inter-participant OR AB intraparticipant OR AB intra-participant OR AB kappa OR AB kappa’s OR AB kappas OR AB repeatab\* OR ((TI replicab\* OR AB replicab\* OR TI repeated OR AB repeated) AND (TI measure OR AB measure OR TI measures OR AB measures OR TI findings OR AB findings OR TI result OR AB result OR TI results OR AB results OR TI test OR AB test OR TI tests OR AB tests)) OR TI generaliza\* OR TI generalisa\* OR TI concordance OR AB generaliza\* OR AB generalisa\* OR AB concordance OR (TI intraclass OR AB intraclass AND TI correlation\* or AB correlation\*) OR TI discriminative OR TI “known group” OR TI factor analysis OR TI factor analyses OR TI dimension\* OR TI subscale\* OR AB discriminative OR AB “known group” OR AB factor analysis OR AB factor analyses OR AB dimension\* OR AB subscale\* OR (TI multitrait OR AB multitrait AND TI scaling OR AB scaling AND (TI analysis OR AB analysis OR TI analyses OR AB analyses)) OR TI item discriminant OR TI interscale correlation\* OR TI error OR TI errors OR TI “individual variability” OR AB item discriminant OR AB interscale correlation\* OR AB error OR AB errors OR AB “individual variability” OR (TI variability OR AB variability AND (TI analysis OR AB analysis OR TI values OR AB values)) OR (TI uncertainty OR AB uncertainty AND (TI measurement OR AB measurement OR TI measuring OR AB measuring)) OR TI “standard error of measurement” OR TI sensitiv\* OR TI responsive\* OR AB “standard error of measurement” OR AB sensitiv\* OR AB responsive\* OR ((TI minimal OR TI minimally OR TI clinical OR TI clinically OR AB minimal OR AB minimally OR AB clinical OR AB clinically) AND (TI important OR TI significant OR TI detectable OR AB important OR AB significant OR AB detectable) AND (TI change OR AB change OR TI difference OR AB difference)) OR (TI small\* OR AB small\* AND (TI real OR AB real OR TI detectable OR AB detectable) AND (TI change OR AB change OR TI difference OR AB difference)) OR TI meaningful change OR TI “ceiling effect” OR TI “floor effect” OR TI “Item response model” OR TI IRT OR TI Rasch OR TI “Differential item functioning” OR TI DIF OR TI “computer adaptive testing” OR TI “item bank” OR TI “cross-cultural equivalence” OR TI outcome assessment OR AB meaningful change OR AB “ceiling effect” OR AB “floor effect” OR AB “Item response model” OR AB IRT OR AB Rasch OR AB “Differential item  #4 #1 AND #2 AND #3  **Search Strategy in PsycINFO**  #1 ((TI numeric\* OR AB numeric\*) AND (TI rating OR AB rating) AND (TI scale OR AB scale OR TI score OR AB score)) OR (TI numeric scale OR AB numeric scale) OR nrs OR nprs  #2 DE "BACKACHE" OR TI back pain OR AB back pain OR lbp OR TI back injury OR AB back injury OR DE "LUMBAR vertebrae" OR TI lumbar spine OR AB lumbar spine OR DE "SACROILIAC joint" OR TI sacroiliac pain OR AB sacroiliac pain OR DE "ZYGAPOPHYSEAL joint" OR TI facet pain OR AB facet pain OR DE "SCIATICA" OR TI sciatica OR AB sciatica OR (DE "SPINAL canal -- Stenosis") OR TI lumbar stenosis OR AB lumbar stenosis OR DE "INTERVERTEBRAL disk displacement" OR TI herniated disk OR AB herniated disk OR TI disk pain OR AB disk pain OR DE "SPONDYLOLISTHESIS" OR TI spondylolisthesis OR AB spondylolisthesis  #3 TI psychometr\* OR TI observer variation OR TI reproducib\* OR TI reliab\* OR TI unreliab\* OR TI valid\* OR TI coefficient OR TI homogeneity OR TI homogeneous OR TI “internal consistency” OR AB psychometr\* OR AB observer variation OR AB reproducib\* OR AB reliab\* OR AB unreliab\* OR AB valid\* OR AB coefficient OR AB homogeneity OR AB homogeneous OR AB “internal consistency” OR (TI cronbach\* OR AB cronbach\* AND (TI alpha OR AB alpha OR TI alphas OR AB alphas)) OR (TI item OR AB item AND (TI correlation\* OR AB correlation\* OR TI selection\* OR AB selection\* OR TI reduction\* OR AB reduction\*)) OR TI agreement OR TI precision OR TI imprecision OR TI “precise values” OR TI test-retest OR AB agreement OR AB precision OR AB imprecision OR AB “precise values” OR AB test-retest OR (TI test OR AB test AND TI retest OR AB retest) OR (TI reliab\* OR AB reliab\* AND (TI test OR AB test OR TI retest or AB retest)) OR TI stability OR TI interrater OR TI inter-rater OR TI intrarater OR TI intra-rater OR TI intertester OR TI inter-tester OR TI intratester OR TI intra-tester OR TI interobserver OR TI inter-observer OR TI intraobserver OR TI intra-observer OR TI intertechnician OR TI inter-technician OR TI intratechnician OR TI intra-technician OR TI interexaminer OR TI inter-examiner OR TI intraexaminer OR TI intra-examiner OR TI interassay OR TI inter-assay OR TI intraassay OR TI intra-assay OR TI interindividual OR TI inter-individual OR TI intraindividual OR TI intra-individual OR TI interparticipant OR TI inter-participant OR TI intraparticipant OR TI intra-participant OR TI kappa OR TI kappa’s OR TI kappas OR TI repeatab\* OR AB stability OR AB interrater OR AB inter-rater OR AB intrarater OR AB intra-rater OR AB intertester OR AB inter-tester OR AB intratester OR AB intra-tester OR AB interobserver OR AB inter-observer OR AB intraobserver OR AB intra-observer OR AB intertechnician OR AB inter-technician OR AB intratechnician OR AB intra-technician OR AB interexaminer OR AB inter-examiner OR AB intraexaminer OR AB intra-examiner OR AB interassay OR AB inter-assay OR AB intraassay OR AB intra-assay OR AB interindividual OR AB inter-individual OR AB intraindividual OR AB intra-individual OR AB interparticipant OR AB inter-participant OR AB intraparticipant OR AB intra-participant OR AB kappa OR AB kappa’s OR AB kappas OR AB repeatab\* OR ((TI replicab\* OR AB replicab\* OR TI repeated OR AB repeated) AND (TI measure OR AB measure OR TI measures OR AB measures OR TI findings OR AB findings OR TI result OR AB result OR TI results OR AB results OR TI test OR AB test OR TI tests OR AB tests)) OR TI generaliza\* OR TI generalisa\* OR TI concordance OR AB generaliza\* OR AB generalisa\* OR AB concordance OR (TI intraclass OR AB intraclass AND TI correlation\* or AB correlation\*) OR TI discriminative OR TI “known group” OR TI factor analysis OR TI factor analyses OR TI dimension\* OR TI subscale\* OR AB discriminative OR AB “known group” OR AB factor analysis OR AB factor analyses OR AB dimension\* OR AB subscale\* OR (TI multitrait OR AB multitrait AND TI scaling OR AB scaling AND (TI analysis OR AB analysis OR TI analyses OR AB analyses)) OR TI item discriminant OR TI interscale correlation\* OR TI error OR TI errors OR TI “individual variability” OR AB item discriminant OR AB interscale correlation\* OR AB error OR AB errors OR AB “individual variability” OR (TI variability OR AB variability AND (TI analysis OR AB analysis OR TI values OR AB values)) OR (TI uncertainty OR AB uncertainty AND (TI measurement OR AB measurement OR TI measuring OR AB measuring)) OR TI “standard error of measurement” OR TI sensitiv\* OR TI responsive\* OR AB “standard error of measurement” OR AB sensitiv\* OR AB responsive\* OR ((TI minimal OR TI minimally OR TI clinical OR TI clinically OR AB minimal OR AB minimally OR AB clinical OR AB clinically) AND (TI important OR TI significant OR TI detectable OR AB important OR AB significant OR AB detectable) AND (TI change OR AB change OR TI difference OR AB difference)) OR (TI small\* OR AB small\* AND (TI real OR AB real OR TI detectable OR AB detectable) AND (TI change OR AB change OR TI difference OR AB difference)) OR TI meaningful change OR TI “ceiling effect” OR TI “floor effect” OR TI “Item response model” OR TI IRT OR TI Rasch OR TI “Differential item functioning” OR TI DIF OR TI “computer adaptive testing” OR TI “item bank” OR TI “cross-cultural equivalence” OR TI outcome assessment OR AB meaningful change OR AB “ceiling effect” OR AB “floor effect” OR AB “Item response model” OR AB IRT OR AB Rasch OR AB “Differential item  #4 #1 AND #2 AND #3 | | | | |  |
|  |  |  |  |  |  |